Zheyuan Brian Zhang

Ann Arbor, MI
zheyuan@umich.edu
cozheyuanzhangde.github.io

Research Interests

Deep Learning, Natural Language Processing, Multimodal Learning, Continual/Online/Few-shot Learning, Robotics and Embodied Intelligence, Cognitive Science (Cognitive Psychology and Neuroscience)

Education

- **University of Massachusetts Amherst** — Amherst, MA, United States
  Computer Science (B.S.), GPA: 3.848/4.000 — *September 2018 - May 2022*
  Computer Science Courses: *Introduction to Problem Solving with Computers, Programming with Data Structures, Programming Methodology, Computer Systems Principles, Reasoning Under Uncertainty, Introduction to Computation, Introduction to Algorithms, Web Programming, Social Issues in Computing (writing), Introduction to Computer Vision, Artificial Intelligence, Introduction to Robotics: Perception, Mechanics, Dynamics, and Control, Game Programming (graduate-level), Machine Learning (graduate-level), Practice and Applications of Data Management, Introduction to Computer Graphics, Probabilistic Graphical Models (doctoral-level)*

- **University of Michigan Ann Arbor** — Ann Arbor, MI, United States
  Robotics (M.S.) and Cognitive Science (Graduate Certificate), GPA: 4.000/4.000 — *August 2022 - April 2024*
  Related Courses: *Mathematics for Robotics, Robotic Systems Laboratory, Natural Language Processing, Computational Modeling of Cognition, Cross-Disciplinary Perspectives in Cognitive Science, Advanced Topics in Computer Vision, Advanced Robot Operating Systems, Advanced Artificial Intelligence*

Research Experience

- **Lab of Autonomous Robotics and Systems, Fudan University** — Shanghai, China
  Research Intern — *June 2021 - August 2021*
  Mentors: Prof.Huiliang Shang, Ruijiao Li

  - **Robotic Grasping and Machine Vision**: Developed software for an robotic arms (Aubo-i5, UR5) and RGB-D cameras (Chishine3D, Intel RealSense D435), enabling tabletop object manipulation using ROS. Integrated YOLOv4 for object localization and classification, executed hand-eye calibration, and implemented grasping procedures involving FK/IK and motion planning with MoveIt. Improved the template matching algorithm for low-cost, robust object detection and grasp pose estimation, while also eliminating redundant bounding boxes.

    Paper accepted by IEEE INSAI 2021 (Awarded Second Prize). [*code: 67* ★]

- **Situated Language and Embodied Dialogue (SLED) Lab, UMich** — Ann Arbor, United States
  Research Assistant — *January 2023 - Now*
  Mentors: Prof.Joyce Chai, Shane Storks, Peter Yu

  - **Physical Commonsense Reasoning**: Inspired by the dual-process theory in cognitive psychology, which theorizes that human often use intuitive heuristic thinking to guide slower analytic reasoning, I applied this cognitively motivated approach to both fine-tuning and in-context learning for physical commonsense reasoning in language models. The proposed methods greatly improved language models' reasoning coherence. Additionally, this enhancement is a direct result of more faithful attention to relevant language context.

    Paper accepted by EMNLP 2023 Main Conference. [*code: 2* ★]
  - **Video In-context Learning**: The proposed method EILEV enhances VLMs' in-context learning capabilities for egocentric videos without needing massive video datasets. EILEV uses architectural adjustments and training data adaptations to allow the model to process contexts interleaved with video clips and narrations. The results show that EILEV-trained models outperform larger VLMs and exhibit superior generalizations.

    Paper submitted and under review. [*web*] [*code: 35* ★]
  - **Language Emergence**: In progress.

Publications

- **Zheyuan Zhang**, Shane Storks, Fengyuan Hu, Sungryull Sohn, Moontae Lee, Honglak Lee, Joyce Chai. "From Heuristic to Analytic: Cognitively Motivated Strategies for Coherent Physical Commonsense Reasoning". *Empirical Methods in Natural Language Processing, EMNLP, 2023.* [arXiv]
- **Zheyuan Zhang**, Huiliang Shang. "Low-cost Solution for Vision-based Robotic Grasping". *International Conference on Networking Systems of AI, IEEE INSAI, 2021.* DOI: 10.1109/INSAI54028.2021.00022.

Writings and Pre-prints

- Keunwoo Peter Yu, **Zheyuan Zhang**, Fengyuan Hu, Joyce Chai. "Efficient In-Context Learning in Vision-Language Models for Egocentric Videos". *2022.* [arXiv]
- **Zheyuan Zhang**, Yu Zhu, Manu Aatitya Raajan Priyadharshini, Thirumalaesh Ashokkumar. "Bot Lab: Autonomous Ground Vehicle from Low-level Control, SLAM to Planning and Exploration". *2022.* [PDF Online]

Honors and Awards

- **Chancellor's Award ($10,000 per year)** | University of Massachusetts Amherst
- **8 × Dean's List Honors (all semesters)** | University of Massachusetts Amherst
- **Second Prize - Excellent Paper** | IEEE INSAI2021
- **ROS Summer School 2021 Certificate** | ROS Education Foundation China

Teaching Experience

- **EECS 492: Introduction to Artificial Intelligence** Ann Arbor, MI, United States
  Graduate Student Instructor (GSI) *Fall 2023, Winter 2024*
  - **Topics**: Agents and Rationality, Problem, Solving and Search, Uninformed Search, Informed Search, Local Search, Genetic Search, Adversarial Search, Constraint Satisfaction, Logical Agents, Propositional Logic and Inference, First-Order Logic and Inference, Uncertainty, Bayesian Networks, Intro to Learning, Decision Trees and Linear Regression, Neural Networks, Neural Network Architectures and Generative AI, Nonparametric Methods (SVM, k-NN), Decision Analysis, Making Complex Decisions and Approximate Inference, Additional Topics in Complex Decisions (MCTS), Reinforcement Learning, Active Reinforcement Learning, Game Theory
  - **Responsibilities**: Lecture attendance (drop-in office hours and special events), office hours, making homeworks and discussion slides, managing Canvas and Piazza (question-answering), proctoring midterm and final exams.

Online Courses

- DeepLearning.AI Deep Learning Specialization, taught by Andrew Ng
- DeepLearning.AI Generative Adversarial Networks (GANs) Specialization, taught by Sharon Zhou
- Reinforcement Learning Course at UCL, taught by David Silver

Skills

Programming Languages and Tools/Software: *Python, C, C++, JavaScript, Java, VB/VB.NET, C#/C#.NET, NASM Assembly, SQL, Shell Script, HTML/CSS, MATLAB, LaTeX, Linux, Robot Operating System (ROS)*

Guest Lectures and Invited Talks

- **EECS 492 (University of Michigan): "Foundation Models for AI: A general introduction to foundation models and applications to reasoning, planning, manipulation and navigation"**

Attended Interdisciplinary Events

- **11th Annual Marshall M. Weinberg Symposium on Computational Psychiatry. March 25, 2023.**

Languages

- **Chinese** (Native)
- **English** (Fluent)
- **Korean** (Beginning)

Projects (during undergraduate and high school)

- **Deep Reinforcement Learning** *January 2022 - January 2022*
  - Designed a game (environment) including reward setups using Pygame for an agent to explore and make actions.
  - Implemented deep Q-learning.
  - Implemented curiosity-driven exploration.

- **FindUrCourse (*www.findurcourse.com*) – Node.js Web Application** *October 2020 – January 2021*
  - Designed a website for college students to search, rank, and rate courses. FindUrCourse currently has more than 70 courses and 100 comments.
  - Implemented front-end user interface using HTML/CSS, JavaScript with Bootstrap framework and back-end server using Node.js with Express framework.
  - Created PostgreSQL database for managing data of users and courses. Implemented server APIs which executes SQL commands on Node.js by pg-promise. Constructed 25 API end-points for front-end functionalities.
  - Implemented hash encryption on the server-side to protect users' passwords security.

- **ZiZoyaOS - 32-Bit Operating System** *December 2020 – January 2021*
  - Developed a bootloader that reads disk sectors, switches the operating system into 32-bit protected mode, loads the kernel, and initializes the GDT using NASM Assembly.
  - Implemented a kernel using C, which has the following functionalities and components: IDT, ISR, IRQ initialization, ports communication, CPU interrupts handler for capturing, processing and giving feedback, display driver, keyboard driver (allows key combinations).
  - Designed a command-line interface and some functionalities, which include a calculator allowing simple arithmetic operations.

- **Quanin - Automatic Stock Screener** *March 2021*
  - Developed a stock screener based on stock trader's demand using Python and various libraries, including Baostock, Pandas, TA-Lib, Matplotlib.
  - Created a graphical user-friendly interface using C Sharp.NET and Winform.

- **Clara in Wonderland - Unity3D Open-world Adventure Game** *November 2021*
  - Designed terrains and environment of an open-world scene including items, enemies, obstacles, etc.
  - Designed a game story using triggered narration and user interface including the main menu and in-game menu.
  - Applied various sound effects including background music, hitting, drinking, eating, picking up, etc.
  - Implemented health bar, hunger bar and thirst bar for elements of survival.
  - Implemented an inventory with interactions and items management with object-oriented programming.
  - Implemented game features. For example, filling water into an empty jar and the character can drink it when thirsty.
  - Developed enemies attacking logic with animator controller and pathfinding using NavMesh.

- **LazyLaTeX for Homeworks - LaTeX Code Generator** *November 2021*
  - Developed a LaTeX code generator which takes in simple commands and generate LaTeX code automatically.
  - Implemented commands myname(), coursename(), homeworkname(), question(), answer(), code().
  - Implemented functionalities including parsing from matrix string to LaTeX matrix.
  - Created a graphical user-interface including quick view using C Sharp.NET and Winform.

- **Other Projects** *2016 - 2021*
  - Implemented a blockchain using C++ and Cryptography Library.
  - Developed several Windows applications using C Sharp.NET and VB.NET.
  - Experienced compiler design by trying to create a Chinese interpreter (not developed) using C, Flex and Bison.

Selected Course Projects

- **Machine Learning** *Fall 2021*
  - Project 1: Implemented k-nearest neighbors classifier and decision tree classifier with hyper-parameter tuning on Pima Indians Diabetes Dataset. Derived posterior probability and Maximum Likelihood Estimation (MLE) estimator (N samples independently drawn from a normal distribution with known variance and unknown mean) and Maximum A Posterior (MAP) estimator.
  - Project 2: Constructed bag-of-words model and implemented multinomial naive Bayes model on Women's Clothing E-Commerce Reviews Dataset. Derived Bayes optimal classifier and MLE estimates. Implemented a l2-regularized logistic regression classifier.
  - Project 3: Implemented a support vector machine (SVM) classifier with linear kernel, polynomial kernel and RBF kernel for Breast Cancer Wisconsin (Diagnostic) Dataset. Developed ensemble models including random forest classifier and AdaBoost classifier using Scikit-learn. Implemented cross-validation procedure.

- Project 4: Derived mathematical proofs for ridge regression. Developed ordinary least squares (OLS) linear, lasso and ridge regression model with hyper-parameter tuning. Derived partial derivatives used for backpropagation and implemented a fully connected neural network using PyTorch library for classifying MNIST digits. Developed stacking classifiers using Scikit-learn.
- Project 5: Developed convolutional neural networks for classifying MNIST digits using PyTorch, Implemented singular value decomposition (SVD) to perform optimal k-rank approximation and sub-optimal k-rank approximation on image compression. Implemented principle component analysis (PCA) on MNIST dataset.
- Final Project: Attended Kaggle PetFinder.my - Pawpularity Contest which aims to predict the popularity of shelter pet photos. Developed traditional machine learning models including linear models and ensemble models using Scikit-learn to train and predict on the dataset. Implemented DL-enhanced model including a CNN with a self-created network architecture for images using PyTorch and lasso regression for meta-data. Implemented a 2-head neural network model which includes a CNN for images and a FCNN for meta-data.

- **Probabilistic Graphical Models** *Spring 2022*
  - Project 1: Derived and implemented maximum likelihood learning in the directed model with a probability query answerer. Computed log-likelihood with five-fold cross-validation to assess the performance of the model. Designed a network structure for the heart disease domain and repeated the assessment.
  - Project 2: Implemented simple exhaustive inference algorithm for the conditional random field (CRF) model. Implemented sum-product message passing inference algorithm. Derived maximum likelihood learning for conditional random field models.
  - Project 3: Implemented average log conditional likelihood as the objective function and its gradient functions using sum-product message passing. Implemented the learning algorithm for optimization of maximizing the objective function using L-BFGS-B algorithm. Applied log-sum-exp trick to resolve numerical underflow and overflow.
  - Project 4: Derived and implemented Gibbs sampling Markov chain Monte Carlo (MCMC) algorithm to perform image denoising using a grid-structured CRF model. Tuned hyper-parameters by grid search.
  - Project 5: Derived evidence lower bound (ELBO) with reparameterization trick. Implemented variational autoencoder (VAE) training algorithm on images of 0/1 picked from the MNIST dataset with the automatic differentiation tool JAX. Applied VAE to dimension reduction and image generalization as a generative model. Estimated the log likelihood of training and test set by Monte Carlo samples. Applied VAE to anomaly detection to non 0/1 images.
- **Natural Language Processing** *Fall 2022*
  - Project 1: Derived mathematical formulas of cross entropy and mutual information. Implemented minimum edit distance using dynamic programming. Implemented regular expression for identifying named entities. Implemented a Naive Bayes Classifier for sentiment analysis.
  - Project 2: Implemented the count-based word vector and prediction-based word vector. Implemented a part-of-speech (POS) tagger using a recurrent neural network (RNN).
  - Project 3: Answered questions regarding probabilistic context free grammar (PCFG) and dependency parsing. Fine-tuned a pre-trained language model to perform question answering using HuggingFace transformers.
  - Project 4: Reproduced experimental results of paper "INFOTABS: Inference on Tables as Semi-structured Data".
- **Robotic Systems Laboratory** *Fall 2022*
  - Arm Lab (Group Project): We calibrated intrinsic camera and automatic workspace calibration using Apriltags and SolvePnP. Implemented a block detector that is capable of detecting colorful blocks and determining their location and poses in the workspace. Implemented forward kinematics and inverse kinematics using DH-table. Implemented state-machine for teach & repeat and competition tasks.
  - Bot Lab (Group Project): We implemented low-level control that executes commands from high-level system to drive the robot based on velocity models and kinematics with a PID controller. Implemented Simultaneous Localization and Mapping (SLAM) from scratch which includes mapping module, particle filter with action model and sensor model. Implemented AStar (A*) heuristic search with pruning algorithm for path planning and frontier-guided algorithm for exploration.